بسم الله الرحمن الرحیم

آمار حیاتی

رفرنس :

۱. آمار حیاتی دکتر کاظم محمد
۲. آمار زیستی دکتر محمد تقی آیت اللهی

مدرس : کاربرد

فصل چهارم

معیارهای گرایش به مرکز

مشخص کننده های عددی

مشخص کننده های عددی

شاخصهای

- **تغییرات**
  - ضریب کجی
  - ضریب گشیدگی
- **پراکندگی**
  - دامنه تغییرات
  - قدر مطلق احرافات
  - واریانس
  - انحراف معیار
  - ضریب تغییرات
- **گرایش مرکزی**
  - میانگین
  - میانه
  - مد

شاخصهای گرایش مرکزی(به انضمام اصطلاح لاتینی آنها)

میانگین هارمونیک(Weighted Mean)　میانگین حسابی (Arithmetic Mean)

میانگین هندسی (Geometric Mean)　مد(Mode)　میانه(Median)

شاخصهای پراگندگی

دامنه تغیرات (Range)

واریانس(Variance)　انحراف معیار (Standard Deviation)

شاخصهای تغییرات

ضریب تغییرات

ضریب گشیدگی(Kurtosis)　ضریب کجی (Skewness)

معیارهای گرایش به مرکز

در آمار همواره سعی می‌شود تا اطلاعات نهفته در کل داده‌ها را به صورت یک یا چند عدد معقول در آورد این اعداد معیارهای گرایش به مرکز نامیده می‌شود. از مهم‌ترین معیارهای گرایش به مرکز عبارتند از :

الف) میانگین

ب) نما

ج) میانه

الف) میانگین :

اصلی‌ترین و مورد استفاده‌ترین معیار گرایش به مرکز میانگین است .

فرض کنید n داده به صورت $x_1 , x_2 ,..., x_k$ با فراوانی‌ها $f_1, f_2 ,..., f_k$ باشند در این صورت

$$\bar{x} = \frac{\text{تعداد کل داده‌ها}}{\text{مجموع کل داده‌ها}} = \frac{\sum_{i=1}^{k} f_i x_i}{n}$$

را میانگین داده‌ها می‌گویند

مثال : میانگین اعداد ۱۱ ، ۹ ،۷ ، ۵ ، ۳ ،۱

$$\bar{x} = \frac{1+3+5+7+9+11}{6} = 6$$

مثال : میانگین حقوق کارمندان اداره‌ای بر حسب واحد ده‌هزار تومان
در جدول فراوانی زیر عبارت است از :

حقوق	f_i	x_i	$f_i x_i$	$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{2360}{40} = 59$
۲۰-۴۰	۵	30	150	
۴۰-۶۰	۱۷	50	850	
۶۰-۸۰	۱۳	70	910	
۸۰-۱۰۰	۵	90	450	
	n=40		$\sum f_i x_i = 2360$	

مثال : میانگین تعداد فرزندان ۲۰ خانواده در جدول فراوانی زیر عبارت است از :

x_i	f_i	$f_i x_i$
0	5	0
1	3	3
2	4	8
3	6	18
4	2	8
	$n = 20$	37

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i \times x_i}{n} = \frac{37}{20} = 1/85$$

karbord2003@yahoo.com ۹

میانگین جامعه:

هر گاه اطلاعات جمع آوری شده از جامعه ای به حجم **N** باشد در این صورت میانگین بدست آمده میانگین جامعه بوده و با μ نشان می دهیم که در آن :

$$\mu = \frac{\sum_{i=1}^{N} f_i x_i}{N}$$

karbord2003@yahoo.com ۱۰

میانگین وزنی

در برخی مطالعات یا اندازه گیریها ممکن است کمیتهای بدست آمده دارای وزن یا ضریب خاصی باشند در این حالت میانگین وزنی محاسبه می شود . که فرمول آن عبارت است از:

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

که در آن w_i وزن داده i ام (x_i) می باشد

مثال : دانشجویی دریک درس۳ واحدی نمره ۱۴، دردرس ۲ واحدی ۱۲، در درس ۴ واحدی ۱۵ و در ۳ واحدی دیگر ۱۱ گرفته است.
معدل وی ویا به بیان آماری میانگین نمرات وی چقدر است؟

x_i	w_i	$w_i x_i$
14	**3**	**42**
12	**2**	**24**
15	**4**	**60**
11	**3**	**33**
مجموع	$\sum w_i = 12$	$\sum w_i x_i = 159$

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} = \frac{159}{12} = 13/25$$

مثال : در یک شهر سه روزنامه محلی منتشر می شود. ۱۸ در صد خانواده‌های ساکن این شهر به هیچیک از روزنامه ها مشترک نیستند . ۶۱ در صد آنها یکی از روزنامه‌ها ، ۱۷ درصد دو روزنامه و۴ درصد دیگر هر سه روزنامه مشترکند .

متوسط اشتراک خانواده های این شهر به این سه روزنامه چقدر است؟

x_i	w_i	$w_i x_i$
0	**18**	**0**
1	**61**	**61**
2	**17**	**34**
3	**4**	**12**
	$\sum w_i = 100$	$\sum w_i x_i = 107$

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} = \frac{107}{100} = 1/07$$

karbord2003@yahoo.com

میانگین میانگین‌ها

فرض کنید k نمونه با حجم‌های $n_1, n_2, \ldots, n_k$ با میانگین‌های $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ باشند در این صورت میانگین میانگین‌ها عبارت است از :

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i}$$

karbord2003@yahoo.com

مثال : میانگین مزد ۱۰ کارگر ۳۵۰ هزار تومان و میانگین مزد ۱۵ کارگر دیگر ۲۵۰ هزار تومان می‌باشد میانگین مزد این ۲۵ کارگر را حساب کنید .

$$n_1 = 10$$
$$\bar{x}_1 = 350$$
$$n_2 = 15$$
$$\bar{x}_2 = 250$$

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i} = \frac{10 \times 350 + 15 \times 250}{25} = 290$$

karbord2003@yahoo.com

ب) نما :

داده‌ای که فراوانی آن نسبت به سایر داده‌ها بیشتر باشد مد یا نما نامیده می‌شود و با M نشان داده می‌شود .

مثال : نمای داده‌های ۳ و ۱ و ۳ و۱ و۳ و ۳ و ۲ و ۴ و ۵ و ۳ و ۳ و ۱

عبارت است از **M=3**

مثال : جدول زیر تعداد واکسیناسیون انجام شده علیه انواع بیماری‌ها ، در یک سال را نشان می‌دهد مد یا نما عبارت است از :

نوع بیماری	فلج اطفال	سرخک	سل	دیفتری
فراوانی بر حسب هزار	**۱۱**	**۳**	**۸**	**۴**

فلج اطفال = M

karbord2003@yahoo.com

مثال : توزیع فراوانی تعداد افراد خانوار به صورت زیر داده شده است .

تعداد افراد خانوار	۱	۲	۳	۴	۵	۶	۷
فراوانی	۵	۱۰	۲۰	۱۰	۴۰	۲۰	۱۰

در این صورت M=5 چون خانواده‌های ۵ نفری بیشترین فراوانی را دارند

مثال : مد یا نما در مجموعه داده‌های ۲ و ۱ و ۳ و۵ و ۲ و۵ و ۳ و ۱ وجود ندارد زیرا فراوانی تمام داده‌ها یکسان می‌باشد .

نما | چندکها | میانه | میانگین

نمودارهاي آماري:

محاسبه نما براي داده‌هاي پيوسته

کلاس	x_i	f_i	r_i
۳۵/۱_۵۵/۱	۴۵/۱	۴	۰۸/۰
۵۵/۱_۷۵/۱	۶۵/۱	۶	۱۲/۰
۷۵/۱_۹۵/۱	۸۵/۱	۱۲	۲۴/۰
۹۵/۱_۱۵/۲	۰۵/۲	۹	۱۸/۰
۱۵/۲_۳۵/۲	۲۵/۲	۸	۱۶/۰
۳۵/۲_۵۵/۲	۲/۴۵	۶	۱۲/۰
۵۵/۲_۷۵/۲	۶۵/۲	۲	۰۴/۰
۷۵/۲_۹۵/۲	۸۵/۲	۳	۰۶/۰
جمع		۵۰	۰۰/۱

از روي جدول ملاحظه مي‌شود که فراواني رده ۷۵/۱_۹۵/۱ داراي بيشترين فراواني است بنابراين به عنوان رده نما در نظر مي‌گيريم.

$$M = L_M + \frac{d_1}{d_1 + d_2} \times \omega$$

۲۴/۰-۱۲/۰ ۲۴/۰-۱۸/۰

$$M = 1.75 + \frac{0.12}{0.12 + 0.06} \times 0.2$$

محاسبه نما برای داده‌های طبقه بندی شده :

ابتدا نماینده‌ی طبقه‌ای را که دارای بیشترین فراوانی است و به طبقه‌ی نما معروف است را به عنوان نما اختیار می‌کنند برای دقت بیشتر می‌توان نما را از فرمول

$$M = L_M + \frac{d_1}{d_1+d_2}.w$$

را بدست آورد .

که در آن :

L_M حد پایین طبقه‌ی نما

d_1 اختلاف فراوانی طبقه‌ی نما با طبقه‌ی قبلی

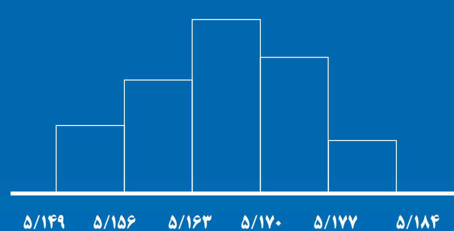
d_2 اختلاف فراوانی طبقه‌ی نما با طبقه بعدی

w طول طبقه

مثال : در جدول فراوانی داده شده مقدار نما را بیابید.

حدود طبقه	فراوانی
149/5–156/5	15
156/5–163/5	2
163/5–170/5	30
170/5–177/5	25
177/5–184/5	10

مقداری که دارای بیشترین فراوانی است بعنوان طبقه نما انتخاب می‌کنیم



طبقه نما یعنی $\frac{163/5+170/5}{2}=167$ بعنوان مقدار تقریبی نما و مقدار دقیق آن عبارت است از :

$$M = L_M + \frac{d_1}{d_1+d_2}.w$$

$$M = 163/5 + \frac{10}{10+5} \times 7$$

$$M = 168/17$$

ج) میانه :

یکی دیگر از معیارهای گرایش به مرکز میانه می‌باشد و آن را با **m** نشان می‌دهند میانه داده‌ای است که تقریباً نصف داده‌ها از آن کوچکتر باشند.

برای بدست آوردن میانه ابتدا داده‌ها را به صورت غیر نزولی از کوچ به بزرگ مرتب می‌کنند اگر تعداد داده‌ها فرد باشد در این صورت داده‌ای که در وسط قرار دارد میانه محسوب می‌شود و اگر تعداد داده‌ها زوج باشد نصف مجموع دو داده‌ای که در وسط قرار دارند میانه محسوب می‌شود .

مثال : برای پیدا کردن میانه داده‌های
۲۰ و ۲۳ و ۲۰ و ۱۹ و ۲۰ و ۲۱ و ۲۷
آنها را به صورت غیر نزولی ۲۷ و ۲۳ و ۲۱ و ۲۰ و ۲۰ و ۲۰و ۱۹
مرتب می‌کنیم لذا میانه عدد وسطی است $m=20$

مثال : میانه داده‌های ۲ و۳ و۵ و۴ و۱ و۲ و۳ و۲ می‌شود $m = 2/5$

اگر داده‌ها گسسته همراه با فراوانی داده شده باشند در این صورت با تشکیل ستون فراوانی تجمعی ، داده‌ای که در وسط قرار دارد را بعنوان میانه انتخاب می‌کنیم.

x_i	۱۲	۱۴	۱۵	۱۷	۱۹
f_i	۴	۸	۶	۴	۳

مثال : جدول توزیع فراوانی نمرات ۲۵ دانشجو به صورت زیر داده شده است .

x_i	f_i	F_i
12	4	4
14	8	12
15	6	18
17	4	22
19	3	25

با تشکیل ستون فراوانی تجمعی و با توجه به $n=24$ ملاحظه می‌شود که داده‌ی وسطی سیزدهمین داده مرتب شده می‌باشد

$M=x_{(13)}=15$

محاسبه میانه برای داده‌های طبقه بندی شده :

ابتدا طبقه‌ای را که نصف داده‌ها در آن قرار دارد و یا به عبارتی اولین طبقه‌ای که فراوانی تجمعی آن برابر نصف یا بیشتر از نصف داده‌ها باشد و به آن طبقه میانه می‌گویند را مشخص کرده و سپس از فرمول زیر میانه را محاسبه می‌کنند.

karbord2003@yahoo.com

نما
چندکها
میانه
میانگین

معیارهاي مرکزي:

محاسبه میانه براي داده‌هاي پيوسته

کلاس	x_i	f_i	F_i
۳۵/۱_۵۵/۱	۴۵/۱	۴	۴
۵۵/۱_۷۵/۱	۶۵/۱	۶	۱۰
۷۵/۱_۹۵/۱	۸۵/۱	۱۲	۲۲
۹۵/۱_۱۵/۲	۰۵/۲	۹	۳۱
۱۵/۲_۳۵/۲	۲۵/۲	۸	۳۹
۳۵/۲_۵۵/۲	۲/۴۵	۶	۴۵
۵۵/۲_۷۵/۲	۶۵/۲	۲	۴۷
۷۵/۲_۹۵/۲	۸۵/۲	۳	۵۰
جمع		۵۰	_

$n/2 = 25$

$$m = L_m + \left(\frac{\frac{n}{2} - F_b}{f_m}\right) w$$

طول هر رده

karbord2003@yahoo.com

$$m = L_m + \frac{n \times 0/5 - F_{m-1}}{f_m} \times w$$

که در آن

n : تعداد کل داده‌ها

L_m : حد پاین طبقه میانه

f_m : فراوانی طبقه میانه

F_{m-1} : فراوانی تجمعی طبقه‌ی ما قبل طبقه میانه

w : طول طبقه

karbord2003@yahoo.com ۲۷

مثال : توزیع فراوانی وزن ۱۰۰ نفر دانشجو در جدول زیر داده شده است میانه وزن دانشجویان عبارت است از :

حدود طبقات	فراوانی	فراوانی تجمعی
59–62	5	5
62–65	18	23
65–68	42	65
68–71	27	92
71–74	8	100
	n = 100	

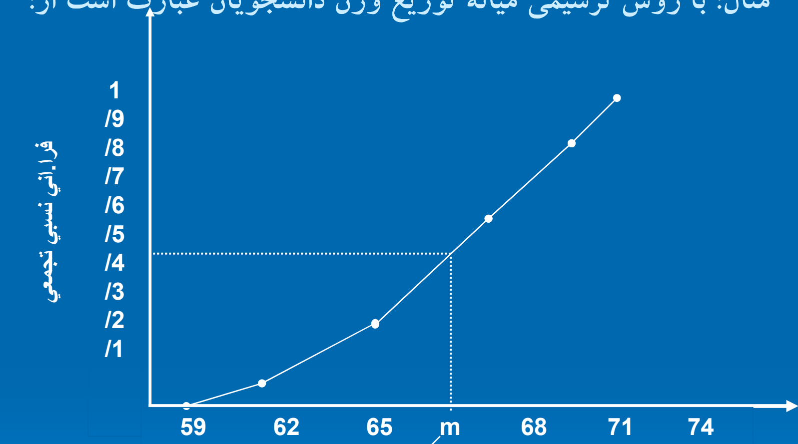
$$m = L_m + \frac{n \times 0/5 - F_{m-1}}{f_m} \times w$$

$$m = 65 + \frac{100 \times 0/5 - 23}{42} \times 3$$

$$m = 66/93$$

karbord2003@yahoo.com ۲۸

مثال: با روش ترسیمی میانه توزیع وزن دانشجویان عبارت است از:



نمودار چند ضلعی نسبی تجمعی وزن دانشجویان

m=66/93

خاصیت مهم میانه این است که حاصل جمع قدر مطلق تفاضلهای مقادیراز میانه، از حاصل جمع قدرمطلق تفاضلهای مقادیر از هر عدد دیگر مثل C، کوچکتر است یعنی

$$\sum_{i=1}^{k} f_i |x_i - m| \leq \sum_{i=1}^{k} f_i |x_i - c|$$

به عبارت دیگر میانه حاصل جمع قدر مطلق تفاضلهای مقادیری از هر عدد دیگر را به حداقل می رساند.

مزایاومعایب میانگین حسابی

مزایا:

۱–به راحتی محاسبه می‌شود

۲–تمام داده هادر محاسبه آن شرکت دارند.

۳–عملیات‌جبری به‌راحتی روی آنها انجام می‌گیرد.

٤–نوسانات نمونه‌گیری حداقل تأثیر راروی آنها داردوازاین جهت به آنها پایا می‌گویند.

معایب:

۱–میانگینهای حسابی تحت تأثیرمقادیر کناره قرار می‌گیرند.

۲–درصورت‌نبودن جزئیات محاسبات میانگین – ها ممکن است نتایج نادرستی را بیان کند.

مزایا ومعایب میانه

مزایا

۱–درک ومحاسبه آن آسان است

۲–اگربتوانیم داده‌هارابه‌صورت‌صعودی یا نزولی مرتب کنیم‌تنها با اندازه گیری یک یا دوتای آنها قابل محاسبه است.

۳–تحت‌تأثیرمقادیرکناری قرارنمی‌گیرد.

٤–برای توزیع‌هایی که‌طبقات کناری باز است میانه یکی ازمشخص‌کننده‌های مرکزی است

معایب

۱–وقتی تعداد مشاهدات زوج است میانه به طوردقیق معلوم نمی‌شود.

۲–تمام افراددرساختمان میانه شرکت ندارند.

۳–عملیات جبری نمی‌توان روی میانه انجام داد.

٤–نوسانات نمونه گیری تأثیر زیادی روی میانه ها دلرند.

مزایا ومعایب نما یا مد

مزایا :

۱-محاسبه آن آسان است

۲-تحت‌تأثیرمقادیرکناری قرارنمی‌گیرد.

۳-حتی‌وقتی‌که‌عرض‌طبقات نامساوی است قابل محاسبه است.

٤-یکی ازمشخص‌کننده‌های‌خوب در پیش‌بینی‌های مؤسسات بازرگانی‌است.

معایب :

۱-درمواردی که توزیع چند نمایی است نمی‌توان یک نمای مشخص برای آن درنظر گرفت.

۲-همه مشاهدات در ساختمان نما دخالت ندارد.

۳-عملیات جبری روی آنها صورت نمی گیرد

٤- بیش ازمیانگین ونما تحت تأثیر نوسانات نمونه قرار می‌گیرند

سوال

مثال: جدول زیر ۷۹۹ نفر را که از نظر سن گروه‌بندی شده‌اند نشان می‌دهد. میانگین، نما، میانه و چارک اول و سوم را محاسبه کنید و به سئوالات زیر پاسخ دهید.

سن	۲۰-۲۵	۲۵-۳۰	۳۰-۳۵	۳۵-۴۰	۴۰-۴۵	۴۵-۵۰	۵۰-۵۵	۵۵-۶۰
نفر	۵۰	۷۰	۱۰۰	۱۸۰	۱۵۰	۱۲۰	۷۰	۵۹

شماره گروه	گروه سنی	فراوانی	فراوانی تجمعی	حد متوسط	طول کلاس
۱	۲۰-۲۵	۵۰	۵۰	۲۲/۵	۵
۲	۲۵-۳۰	۷۰	۱۲۰	۲۷/۵	۵
۳	۳۰-۳۵	۱۰۰	۲۲۰	۳۲/۵	۵
۴	۳۵-۴۰	۱۸۰	۴۰۰	۳۷/۵	۵
۵	۴۰-۴۵	۱۵۰	۵۵۰	۴۲/۵	۵
۶	۴۵-۵۰	۱۲۰	۶۷۰	۴۷/۵	۵
۷	۵۰-۵۵	۷۰	۷۴۰	۵۲/۵	۵
۸	۵۵-۶۰	۵۹	۷۹۹	۵۷/۵	۵
		۷۹۹			

فصل پنجم

معیارهای پراکندگی

داده‌ها را معمولاً به صورت یک عدد به نام معیار تمرکز خلاصه می‌کنند وقسمتی از اطلاعات موجود در آنها را در این عدد منعکس می‌سازند ولی در تحلیلهای‌آماری ، معیارهای مرکزی به تنهایی نمی‌توانند میزان تنوع ، تفاوت و یا به عبارتی دیگر میزان پراکندگی داده‌ها را بیان‌کنند.

فرض کنید دو مجموعه داده **B** و**A** عبارتند از :

A= ۱ و ۲ و ۳ و ۴ و۴ و۶ و ۸

B =۰ و ۱ و ۲ و ۴ و ۴ و۸ و ۹

ملاحظه می شود که میانگین ،میانه و نمای هر دو مجموعه ، عدد ۴ می باشددر حالی که میزان تجمع و پراکندگی دو مجموعه با هم متفاوت است .

بنابراین معرفی معیارهایی که بتواند بیانگر میزان اختلافها ، تنوعها و میزان پراکندگی مجموعه داده‌ها باشد ضرورت دارد.اینک به معرفی چند معیار مهم برای سنجش پراکندگی می‌پردازیم فرض کنید x_k و... و x_2 و x_1 یک سری داده n تایی با فـراوانی‌های f_k و...وf_2وf_1 با میانگین $\bar{x}$ باشند .

$$R = \max(x_i) - \min(x_i) = x_{(n)} - x_{(1)}$$

الف) دامنه تغییرات

تفاضل کوچکترین داده از بزرگترین داده را که با **R** نشان می‌دهند دامنه تغییرات می گویند .

مثال : حقوق پرداختی ماهیانه کارکنان یک شرکت کوچک به شرح زیر است .

۶۰ و ۲۲ و ۲۰ و ۱۷ و ۱۷ و ۱۷ و ۱۶ و ۱۵

دامنه تغییرات **R = ۶۰ – ۱۵ = ۴۵** می باشد .

ب) میانگین قدر مطلق انحرافات :

معیاری است که بتواند ، معیار مناسب در سنجش پراکندگی داده‌ها میزان انحراف و پراکندگی کل داده‌ها را در مجموعه داده‌ها اندازه بگیرد . پراکندگی و انحراف زمانی مفهوم پیدا می کند که داده ها نسبت به یک مبدأ یا مرکز مقایسه شوند مناسب ترین مرکز برای داده‌ها ، میانگین آنهاست . بنابراین تعریف می کنیم :

$$\text{میانگین قدر مطلق انحرافات} = \frac{\text{مجموع کل قدر مطلق انحرافات}}{\text{تعداد کل انحرافات}}$$

$$M.A.D = \frac{\sum_{i=۱}^{n} |x_i - \bar{x}|}{n}$$

karbord2003@yahoo.com

مثال : میزان مهارت مدیران یک سازمان در برنامه ریزی به صورت زیر است.

x_i = ۸۰ و ۸۵ و ۶۰ و ۶۵ و ۵۰ و ۳۵ و ۴۵

$$\bar{x} = \frac{\sum x_i}{n} = \frac{420}{7} = 6$$

در این صورت

x_i	80	85	60	65	50	35	45	$\sum_{i=1}^{7} x_i$
$\lvert x_i - \bar{x}\rvert$	۲۰	۲۵	۰	5	10	25	15	$\sum_{i=1}^{7} \lvert x_i - \bar{x}\rvert = 100$

$$M.A.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} = \frac{100}{7} = 14/28$$

karbord2003@yahoo.com

محاسبه D.A.M در جدولهای فراوانی

اگر جدول فراوانی دارای K طبقه با نماینده طبقات x_k و... و x_2 و x_1 وفراوانهای متناظر f_k و ... و f_2 و f_1 باشد در این صورت

$$M.A.D = \frac{\sum_{i=1}^{k} f_i \left| x_i - \bar{x} \right|}{n}$$

مثال: در جدول توزیع فراوانی زیر سن ۲۰۰ مرد زیر ۴۰ سال در اولین ازدواج داده شده است.

سن ازدواج	f_i	x_i	$f_i x_i$	$\left\|x_i - \bar{x}\right\|$	$f_i \times \left\|x_i - \bar{x}\right\|$
14/5-19/5	18	17	306	8/9	160/2
19/5-24/5	74	22	1628	3/9	288/6
24/5-29/5	62	27	1674	1/1	682
29/5-34/5	26	32	832	6/1	158/6
34/5-39/5	20	37	740	11/1	222
مجموع	۲۰۰		۵۱۸۰		۱۵۱۱/۴

$$\bar{x} = \frac{\sum f_i x_i}{n} = \frac{5180}{200} = 25/9$$

$$M.A.D = \frac{\sum f_i \left| x_i - \bar{x} \right|}{n} = \frac{1511/4}{200} = 7/557$$

ج) واریانس :

در اندازه‌گیری مجموع انحرافات و پراکندگی از مرکز داده‌ها، انحرافـــات منفی متـأثر از داده‌های کوچکتراز میانگین با انحرافات مثبت خنثی می‌شوند روشی دیگر برای رفع این معضل آن است که انحرافها را مربع کنیم.

karbord2003@yahoo.com

دراین صورت معیـــار میانگین مربعات انحرافـــات از میانگین به عنوان معیار سنـــجش پراکندگی تعریف می شود. و آن را *واریانس* می نامیم .

$$s^{۲} = \frac{\sum_{i=۱}^{n}(x_i - \bar{x})^{۲}}{n}$$

با استفاده از جبر مقدماتی می توان فرمول فوق را به صورت خلاصه تر بازنویسی نمود .

$$s^{۲} = \frac{\sum_{i=۱}^{n} x_i^{۲}}{n} - \bar{x}^{۲}$$

karbord2003@yahoo.com

مثال: سود ده ماه یک شرکت سرم سازی به این صورت است.

۴ و ۶ و ۸ و ۲ و ۳ و ۴ و ۷ و ۵ و ۱ و ۰ = x_i

بنابراین

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{60}{10} = 6$$

ونیز به روش خلاصه تر می توان واریانس را محاسبه نمود .

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \frac{220}{10} - (۴)^۲ = 6$$

انحراف استاندارد

$$s = \sqrt{s^2}$$

جذر مثبت واریانس را انحراف استاندارد می گویند

واریانس جامعه :

$$\delta^2 = \frac{\sum_{i=1}^{N} f_i (x_i - \mu)^2}{N}$$

و گاهـــــی با

$$\delta_x^2 = \frac{\sum_{i=1}^{N} f_i (x_i - \mu_x)^2}{N}$$

نشان می دهند

انحراف استاندارد جامعه را با جذر مثبت از واریانس جامعه می‌توان بدست آورد

محاسبه واریانس در جدولهای فراوانی

اگر جدول فراوانی دارای K طبقه باشد نماینده طبقات را با
x_k و...و x_2 و x_1 با فراوانیهای f_k و ... و f_2 و f_1 نمایش دهند
در این صورت فرمول واریانس را می توان به صورت زیر بازنویسی نمود .

$$s^2 = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^2}{n}$$

و نیز با استفاده از جبر مقدماتی می توان فرمول فوق را خلاصه تر نمود .

$$s^2 = \frac{\sum_{i=1}^{k} f_i x_i^2}{n} - \bar{x}^2$$

karbord2003@yahoo.com ٤٧

مثال:توزیع فراوانی سن ۲۰۰ مرد زیر ۴۰ سال در اولین ازدواج در جدول زیر آمده است .

سن بر حسب سال	f_i	x_i	$f_i x_i$	$f_i x_i^2$
۱۴/۵–۱۹/۵	۱۸	۱۷	۳۰۶	۵۲۰۲
۱۹/۵–۲۴/۵	۷۴	۲۲	۱۶۲۸	۳۵۸۱۶
۲۴/۵–۲۹/۵	۶۲	۲۷	۱۶۷۴	۴۵۱۹۸
۲۹/۵–۳۴/۵	۲۶	۳۲	۸۳۲	۲۶۶۲۴
۳۴/۵–۳۹/۵	۲۰	۳۷	۷۴۰	۲۷۳۸۰
مجموع	۲۰۰		۵۱۸۰	۱۴۰۲۲۰

$$\bar{x} = \frac{\sum_{i=1}^{x} f_i x_i}{n} = \frac{۵۱۸۰}{۲۰۰} = ۲۵/۹$$ سال

$$s^2 = \frac{\sum_{i=1}^{k} f_i x_i^2}{n} - \bar{x}^2 = \frac{۱۴۰۲۲۰}{۲۰۰} - (۲۵/۹)^2 = ۳۰/۲۹$$

$$s = \sqrt{۳۰/۲۹} = ۵/۵$$ سال

karbord2003@yahoo.com ٤٨

هـ)ضریب تغییرات :

نسبت انحراف استاندارد به میانگین یعنی: $C.V = \frac{s}{\bar{x}}$

را که اغلب به صورت درصد بیان می‌شود را ضریب تغییرات

می‌نامند ، این معیار که به واحد اندازه‌گیری بستگی ندارد ؛ در

عمل برای مقایسه بکار می رود.

karbord2003@yahoo.com ٤٩

مثال : کارخانه ای دو نوع لاستیک تولید می کند. برای نوع A میانگین عمر ۱۰۰۰ کیلومتر با انحراف استاندارد ۲۰۰ کیلومتر برای نوع B میانگین عمر ۱۱۰۰ کیلومتر با انحراف استاندارد ۱۰۰۰کیلومتر می باشد،

کدام نوع لاستیک بهتر است؟

$$CV_A = \frac{S_A}{\bar{X}_A} = \frac{200}{1000} = 0/2 \qquad C.V_B = \frac{S_B}{\bar{X}_B} = \frac{1000}{1100} = 0/91$$

لاستیک نوع A بهتر است زیرا دارای ضریب تغییرات کمتری است.

karbord2003@yahoo.com ٥٠

karbord2003@yahoo.com ٥١